Hirofumi Katsuno & Alberto O. Mendelzon: „On the Difference between Updating a Knowledge Base and Revising it“

Universität Hamburg
FB Informatik
Projektseminar:
Projekt Robuste
Verarbeitung von
Instruktionen, WiSe
05/06
Carola Eschenbach,
Christopher Habel,
Özgür Özçep,
15. November 2005
Myriam Lipprandt

- **Einleitung**
  - Beispiel zur Beschreibung verschiedener Modelle und Distanzen
  - Unterscheidung zwischen Revision und Update
- **Vorbereitungen/Grundlagen**
  - Revision und die AGM Postulate
  - Ordnungen zwischen Interpretationen
- **Update**
  - „Possible Models Approach“
  - Update Postulate
- **Schließen mit Aktionen**
- **Kontraktion und „Erasure“**
- **Zeit, Revision und Update**
- **Diskussion**

Einleitung

- Modifikation der KB: es werden zwei Arten unterschieden
  1. Update: Die Welt hat sich geändert; die KB muss auf den neuesten Stand gebracht werden (dynamische Weltsicht).
     - z.B. die Aktion eines Roboters hat die Welt verändert; diese Änderung muss in die KB inkorporiert werden
     - Die AGM-Postulate müssen geändert werden, um Update zu beschreiben
  2. Revision: Neue Informationen über schon vorher betrachtete Situationen (statische Weltsicht)
     - AGM beschreibt nur Revision

Einleitung

- Die KB wird durch Theorien ψ, die in aussagenlogischer Form repräsentiert sind, dargestellt.
- Die neue Information wird durch das Zeichen μ dargestellt.
- Es gibt jeweils verschiedene Modelle der KB ψ und der neuen Information μ
  - Angenommen, die KB ψ wird revidiert mit dem Satz μ.
  - Wenn die Revisionsmethode die AGM-Postulate erfüllt, wählt sie diejenigen Modelle von μ aus die besonders „nahe“ zu den Modellen der KB sind
  - „Nähe“ wird nach einer geordneten Relation bestimmt
  - Die ausgewählten Modelle bestimmen die Menge der neuen KB.
    - ➔ $\psi \circ \mu$

Beispiel zur Beschreibung verschiedener Modelle und Distanzen

KB beschreibt fünf Objekte in einem Raum: A, B, C, D, E.

In dem Raum befindet sich ein Tisch. Die Objekte liegen entweder auf dem Tisch oder sie liegen dort nicht.

Die neue Information (Satz) *a sagt: „Objekt A ist auf dem Tisch“* und gilt auch für *b, c, d, e*.

Die KB ψ besteht aus den Sätzen:

$$(a \wedge \neg b \wedge \neg c \wedge \neg d \wedge \neg e) \vee (\neg a \wedge \neg b \wedge c \wedge d \wedge e)$$

D.h. entweder das Objekt A ist auf dem Tisch oder die Objekte C, D und E sind es.

➔ **Diese KB hat genau zwei Modelle *I* und *J***

Beispiel zur Beschreibung verschiedener Modelle und Distanzen

- Ein Roboter hat die Aufgabe, diese Welt zu verändern. Er soll eine Situation schaffen, in der entweder alle Objekte auf dem Tisch liegen oder keine.
- Diese Änderung kann mit der Inkorporation der neuen Sätze μ modelliert werden.

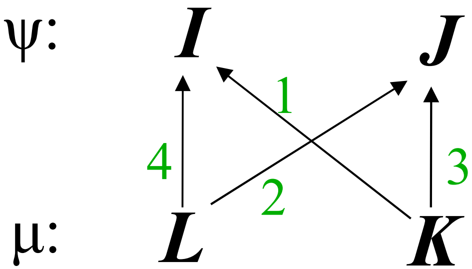
$$(a \wedge b \wedge c \wedge d \wedge e) \vee (\neg a \wedge \neg b \wedge \neg c \wedge \neg d \wedge \neg e)$$

- Die Modelle von μ bekommen die Bezeichnung ***K*** und ***L***
- Die Nähe zwischen Modellen kann durch das Abstandsmaß von Dalal bestimmt werden.
- Die Distanz zwischen Modellen ist die Anzahl der Buchstaben, in denen sich zwei Modelle unterscheiden.

- **Die für die neue KB ausgewählten Modelle sind diejenigen Modelle aus μ mit der minimalsten Distanz zu Modellen von ψ**

Unterscheidung zwischen Revision und Update

Die Distanzen der Modelle sind folgende:



Nach Dalals Revisionsoperator würde man nur das Modell ***K*** auswählen.

- Unser Wissen: Entweder sind alle Objekte auf dem Tisch oder kein Objekt ist auf dem Tisch.
- Es gibt keinen Grund zu glauben, dass nur ein Modell gibt!!!
- Rationales Verhalten bei Revision; in den neuen möglichen Welten müssen alle Modelle von μ wahr sein, daher kann man einige der alten Welten „vergessen“: Sie sind zu unterschiedlich zu der neuen Information.

Unterscheidung zwischen Revision und Update

- Für eine sich ändernde Welt ist der Revisionsoperator ungeeignet. Es existieren mögliche Welten in ψ, von denen wir nicht wissen, welches die „echte“ Welt ist. Wenn die „echte“ Welt sich ändert, versuchen wir alle alten Welten zu untersuchen, um den minimalen Weg der Änderung zu finden, damit wir ein Modell von μ bekommen
    - Die Änderung der Welt gibt uns keinen Grund zu glauben, dass einige der alten Welten nicht möglich sein konnten oder falsch waren.
- Ein Update-Postulat sollte jeder der möglichen alten Welten gleiche Beachtung schenken.

Vorbereitungen

- Aussagenlogische Sprache ***L*** besteht aus Konjunktion von propositionalen Variablen.
- KB wird dargestellt durch aussagenlogische Formeln ψ
- Interpretation von ***L*** ist eine Abbildung der propositionalen Variablen auf {T.F}
- Mod(ψ) beschreibt die Menge aller Modelle von ψ

Revision und die AGM Postulate

- Der Revisionsoperator * auf einer Wissensmenge induziert einen Revisionsoperator $\circ$ auf aussagenlogischen Formeln ψ (Formelmengenrevision)
- (RF 1) $\psi \circ \mu$ impliziert μ.
- (RF 2) Wenn $\Psi \wedge \mu$ erfüllbar ist, dann ist $\psi \circ \mu \leftrightarrow \Psi \wedge \mu$
- (RF 3) Wenn μ erfüllbar ist, dann ist auch $\psi \circ \mu$ erfüllbar.
- (RF 4) Wenn $\psi_1 \leftrightarrow \psi_2$ und $\mu_1 \leftrightarrow \mu_2$, dann auch $\psi_1 \circ \mu_1 \leftrightarrow \psi_2 \circ \mu_2$.
- (RF 5) $(\psi \circ \mu) \wedge \phi$ impliziert $\psi \circ (\mu \wedge \phi)$
- (RF 6) Wenn $(\psi \circ \mu) \wedge \phi$ erfüllbar ist, so gilt: $\psi \circ (\mu \wedge \phi)$ impliziert $(\psi \circ \mu) \wedge \phi$

Ordnungen zwischen Interpretationen

Charakterisierung minimaler Änderungen

- $\boldsymbol{\Omega}$**:** Menge aller Interpretationen von $\boldsymbol{L}$
- Halbordnung $\leq$ über $\boldsymbol{\Omega}$ ist eine reflexive und transitive Relation auf $\boldsymbol{L}$
- $<$ wird Definiert als $\boldsymbol{I} < \boldsymbol{I'}$ gdw. $\boldsymbol{I} \leq \boldsymbol{I'}$ und $\boldsymbol{I'} \not\leq \boldsymbol{I}$

- Die Halbordnung ist total wenn, für jede $\boldsymbol{I}, \boldsymbol{J}$ entweder $\boldsymbol{I} \leq \boldsymbol{J}$ oder $\boldsymbol{J} \leq \boldsymbol{I}$ ist.
- Eine Funktion die jeder Formel ψ eine Halbordnung $\leq$ über $\boldsymbol{\Omega}$ zuweist, nennt man *treu* wenn, sie den folgenden drei Bedingungen genügt:
  1. Wenn $\boldsymbol{I}, \boldsymbol{I'} \in \mathrm{Mod}(\psi)$ dann $\boldsymbol{I} <_\psi \boldsymbol{I'}$ gilt nicht
  2. wenn $\boldsymbol{I} \in \mathrm{Mod}(\psi)$ und $\boldsymbol{I'}$ nicht-$\in \mathrm{Mod}(\psi)$ dann $\boldsymbol{I} <_\psi \boldsymbol{I'}$ gilt
  3. wenn $\psi \leftrightarrow \phi$, dann $\leq_\psi = \leq_\phi$

Minimale Modelle

- $\boldsymbol{M}$ ist eine Teilmenge von $\boldsymbol{\Omega}$.
- Eine Interpretation $\boldsymbol{I}$ ist minimal in $\boldsymbol{M}$ wenn, $\boldsymbol{I} \in \boldsymbol{M}$ und es kein $\boldsymbol{I'} \in \boldsymbol{M}$ so das $\boldsymbol{I'} <_{\psi} \boldsymbol{I}$ gibt.
- Min($\boldsymbol{M}, \leq_{\psi}$) ist die Menge aller Interpretationen in $\boldsymbol{M}$ so das $\boldsymbol{I}$ minimal in $\boldsymbol{M}$ in Bezug auf die Relation $\leq_{\psi}$ ist.
  - **Theorem 1**: Der Revisionsoperator erfüllt die Bedingungen R1 bis R6 wenn, eine *treue* Anweisung existiert die jede KB ψ auf eine totale Halbordnung $\leq_{\psi}$, wie $\mathrm{Mod}(\psi \circ \mu) = \mathrm{Min}(\mathrm{Mod}(\mu), \leq_{\psi})$ abbildet

Update
„Possible Models Approach“

- PMA ist ein Updateoperator da, er jede der möglichen Welten unabhängig voneinander unterstützt.

$$Mod(\psi \diamond_{pma} \mu) = \bigcup_{I \in Mod(\psi)} \text{Incorporate}(Mod(\mu), I),$$

- Incorporate(Mod(μ), ***I***) ist die Menge der Modelle aus μ, die am nächsten an ***I*** „dran“ sind.
- Die Modelle der neuen KB ($\Psi \diamond_{pma} \mu$) ist die Vereinigung der ausgewählten Modelle.
- *Nähe* zwischen zwei Interpretationen ***I*** und ***J*** wird anhand von propositionalen Variablen, die unterschiedliche Wahrheitswerte haben, gemessen: Diff(***I, J***). Für zwei Interpretationen J_1 und J_2 gilt, dass J_1 näher an ***I*** ist als J_2, wenn Diff (I, J_1) eine Teilmenge von Diff(I, J_2) ist. Dann gilt Min(Mod(μ), $\leq_{pma}$).

Beispiel

- L besteht aus zwei propositionalen Variablen: b und m
- KB Ψ : $(b \wedge \neg m) \vee (\neg b \wedge m)$
- $\mu \leftrightarrow b$
- Interpretation ***I*= **{F, T} ist ein Modell von Ψ
- $\boldsymbol{J_1}$= {T,T} und $\mathbf{J_2}$= {T,F} sind zwei Modelle von μ
- $\boldsymbol{J_1} \leq_{\boldsymbol{I},\mathrm{pma}} \boldsymbol{J_2}$ folgt aus dem Fakt das Diff(***I*, *J*₁**)**=**{b} eine Teilmenge von Diff(***I*, *J*₂**)**=**{b, m}
- Interpretation des Beispiel
  - b = Buch, m=Magazin
  - Ψ : entweder das Buch oder das Magazin liegt auf dem Boden aber nicht beides.
  - Roboter soll Buch auf den Boden legen

Beispiel

- Roboter soll Buch auf den Boden legen
- Unser Wissen: Buch liegt auf dem Boden
- Auch Resultat des PMA
- Revisions Postulat R2 Ergebnis: $b \wedge \neg m$
- Wieso sollen wir davon ausgehen, das Magazin liegt nicht auf dem Boden?

Update-Postulate

(U1) $\psi \diamond \mu$ implies μ.

(U2) If ψ implies μ then $\psi \diamond \mu$ is equivalent to ψ.

(U3) If both ψ and μ are satisfiable then $\psi \diamond \mu$ is also satisfiable.

(U4) If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$ then $\psi_1 \diamond \mu_1 \leftrightarrow \psi_2 \diamond \mu_2$.

(U5) $(\psi \diamond \mu) \wedge \phi$ implies $\psi \diamond (\mu \wedge \phi)$.

(U6) If $\psi \diamond \mu_1$ implies μ_2 and $\psi \diamond \mu_2$ implies μ_1 then $\psi \diamond \mu_1 \leftrightarrow \psi \diamond \mu_2$.

(U7) If ψ is complete then $(\psi \diamond \mu_1) \wedge (\psi \diamond \mu_2)$ implies $\psi \diamond (\mu_1 \vee \mu_2)$.

(U8) $(\psi_1 \vee \psi_2) \diamond \mu \leftrightarrow (\psi_1 \diamond \mu) \vee (\psi_2 \diamond \mu)$.

- U1 bis U5 korrespondieren zu R1 bis R5, nur U2 ist schwächer als R2
- Postulat U2: wenn die neue Information μ aus Ψ Ableitbar ist dann, ist bringt Update keine Veränderung in der KB
- Aus dem Postulat R2 folgt **Lemma 3.1**: wenn ein Updateoperator ◊ U2 erfüllt und Ψ ist inkonsistent dann, ist Ψ ◊ μ auch inkonsistent für alle μ.
- Update *repariert* keine KB!

Update-Postulate

- R6 fällt weg, dafür kommen U6-U8 hinzu
- U6: wenn aus $\Psi \lozenge \mu_{1,}\ \ \mu_2$ abgeleitet werden kann und umgekehrt dann, haben sie den gleichen Effekt auf die KB
- U8: *Disjunction rule:* Sie garantiert das jede mögliche Welt der KB unabhängig beachtet wird
- Lemma 2 : Ist $\lozenge$ ein Update-Operator, der (U2) und (U8) erfüllt, so gilt: $\Psi \wedge \mu \models \Psi \lozenge \mu$
- Lemma 3: Ist $\lozenge$ ein Update-Operator, der (U8) erfüllt, so gilt die folgende Monotonie-Eigenschaft: Gilt $\Phi \models \Psi$, dann $\Phi \lozenge \mu \models \Psi \lozenge \mu$.
- Bei den Revisionsoperatoren ist die Monotonie-Eigenschaft nicht gegeben

Update-Postulate

- Formalisierung der Idee von *Nähe* zwischen den Modellen
  - Anstatt jede KB mit einer Ordnung zu versehen, sollte man eine Funktion haben die jede Interpretation auf eine partielle Halbordnung abbildet
  - die Zuweisung ist *Treu* wenn, folgendes gilt:
    - Für beliebige $\boldsymbol{J} \in \Omega$ wenn, $\boldsymbol{I} \neq \boldsymbol{J}$ dann $\boldsymbol{I} <_{\mathrm{I}} \mathbf{J}$
    - Es folgen daraus Theoreme

Theorem 3.4 *Let $\diamond$ be an update operator. The following conditions are equivalent:*

1. *The update operator $\diamond$ satisfies Conditions (U1)$\sim$(U8).*
2. *There exists a faithful assignment that maps each interpretation I to a partial pre-order $\leq_I$ such that*
$$Mod(\psi \diamond \mu) = \bigcup_{I \in Mod(\psi)} Min(Mod(\mu), \leq_I).$$
3. *There exists a persistent assignment that maps each interpretation I to a partial order $\leq_I$ such that*
$$Mod(\psi \diamond \mu) = \bigcup_{I \in Mod(\psi)} Min(Mod(\mu), \leq_I).$$

- Revision definiert eine totale Halbordnung
- Update definiert eine partielle Halbordnung
- Ersetzung der Update-Postulate U6 und U7 durch U9 so ist auch eine totale Halbordnung bei Update möglich

(U9) If ψ is complete and $(\psi \diamond \mu) \wedge \phi$ is satisfiable then $\psi \diamond (\mu \wedge \phi)$ implies $(\psi \diamond \mu) \wedge \phi$.

- Wichtiger Unterschied zwischen Revision und Update: Für Update werden unterschiedliche Ordnungen für jedes Modell der KB eingeleitet. Bei Revision wird nur eine Ordnung durch die ganze KB eingeführt.
- Update hat daher ein lokales Verhalten, Revision ein globales.

Schließen über Aktionen

Eine Aktion kann als ein Paar aus Vor- und Nachbedingung verstanden werden.

- Vorbedingung: wie muss die Welt sein damit eine Aktion ausgeführt werden kann
- Nachbedingung: Beschreibung der Konsequenzen aus der Aktion
- Updateoperator kann zum Schließen über Aktionen benutzt werden
- KB Ψ wird aus α (Vorbed.) und β (Nachbed.) gebildet
- Wenn ψ nicht α impliziert und α ist korrekt und die Aktion ausgeführt werden kann dann gilt: $\psi \Diamond \beta$
- Was passiert wenn, die Vorbedingung nicht erfüllt ist?
  - der Roboter wird Ergebnisse liefern
    - α war richtig oder α war falsch
    - α ist fehlgeschlagen und die Aktion konnte nicht ausgeführt werden
- Bei einem komplizierteren Modell kann z.B. die Aktion nicht ausgeführt werden obwohl α wahr ist oder α kann nicht verifiziert oder falsifiziert werden

Schließen über Aktionen

- Aktion ausgeführt aber α muss revidiert werden: $(\psi \circ \alpha) \Diamond \beta$
- ist α falsch: $\psi \circ -\alpha$
- ist α unbestimmt: Kontraktion!

Kontraktion

- Kontraktion ist der Verlust an Vertrauen bzgl. einiger Sätze

(C1) ψ implies $\psi \bullet \mu$.

(C2) If ψ does not imply μ then $\psi \bullet \mu$ is equivalent to ψ.

(C3) If μ is not a tautology then $\psi \bullet \mu$ does not imply μ.

(C4) If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$ then $\psi_1 \bullet \mu_1 \leftrightarrow \psi_2 \bullet \mu_2$.

(C5) $(\psi \bullet \mu) \wedge \mu$ implies ψ.

Erasure

- Erasure ist zu Kontraktion wie Update zu Revision
- Das Ausradieren von Sätzen einer KB hat das vermehren von Modellen zur Folge.
- Für jedes Modell ***I*** fügt man diejenigen Modelle hinzu die am Nächsten zu ***I*** sind, bei Sätzen in denen μ falsch ist

(E1) ψ implies $\psi \blacklozenge \mu$. st

(E2) If ψ implies $\neg\mu$ then $\psi \blacklozenge \mu$ is equivalent to ψ.

(E3) If ψ is satisfiable and μ is not a tautology then $\psi \blacklozenge \mu$ does not imply μ.

(E4) If $\psi_1 \leftrightarrow \psi_2$ and $\mu_1 \leftrightarrow \mu_2$ then $\psi_1 \blacklozenge \mu_1 \leftrightarrow \psi_2 \blacklozenge \mu_2$.

(E5) $(\psi \blacklozenge \mu) \wedge \mu$ implies ψ.

(E8) $(\psi_1 \vee \psi_2) \blacklozenge \mu$ is equivalent to $(\psi_1 \blacklozenge \mu) \vee (\psi_2 \blacklozenge \mu)$.

Erasure

- Beispiel b = Buch, m = Magazin
  - entweder das Buch liegt auf dem Boden oder das Magazin aber es liegt nicht beides gleichzeitig auf dem Boden
  - Kontrahieren mit b: Ψ ● b ist äquivalent mit Ψ falls die KB b nicht implizierte, d.h Kontraktion ändert nichts wenn, das Buch nicht auf dem Boden liegt
  - Ausradierung hat nur Einfluss auf die Welten in denen das Buch schon auf dem Boden liegt, und es bildet so neue Modelle
    - Ψ ♦ μ ist äquivalent zu $(b \wedge \neg m) \vee \neg b$
    - das neue Modell ist $\neg b \wedge \neg m$
- Kontraktion von b sagt: nichts hat sich geändert aber, falls die KB glaubt das Buch liegt auf dem Boden dann nehme diesen Glauben zurück
- *Erasing b* sagt: hat sich der Zustand des Raumes so geändert das das Buch schon vorher auf dem Boden lag, so ist es jetzt dort nicht mehr

Theorem 5.1

1. *If an update operator* $\diamond$ *satisfies (U1)~(U4) and (U8), then the erasure operator* $\blacklozenge$ *defined by* $(U \to E)$ *satisfies (E1)~(E5) and (E8).*

2. *If an erasure operator* $\blacklozenge$ *satisfies (E1)~(E4) and (E8), then the update operator* $\diamond$ *defined by*

$$\psi \diamond \mu \leftrightarrow (\psi \blacklozenge \neg\mu) \wedge \mu \qquad (E \to U)$$

satisfies (U1)~(U4) and (U8).

3. *Suppose that an update operator* $\diamond$ *satisfies (U1)~(U4) and (U8). Then, we can define an erasure operator by* $(U \to E)$*. The update operator obtained from the erasure operator by* $(E \to U)$ *is equal to the original update operator* $\diamond$*.*

4. *Suppose that an erasure operator* $\diamond$ *satisfies (E1)~(E5) and (E8). Then, we can define an update operator by* $(E \to U)$*. The erasure operator obtained from the update operator by* $(U \to E)$ *is equal to the original erasure operator* $\blacklozenge$*.*

Symmetrisches *Erasure*

- Unvorhersehbare Änderungen, μ und $-\mu$ spielen dieselbe Rolle .
- *Erasure* hatte keine Auswirkung auf μ sondern nur auf $-\mu$
- symmetrisches *Erasure* von b : z.B es besteht kein Wissen über den Ort des Buchs . Das Resultat des symmetrischen *Erasing* ist das es keine Abhängigkeiten mehr bzgl. des Magazin gibt, die KB hat keine Information.

Zeit, Revision und Update

- bisher: keine zeitlichen Parameter, Zeit war implizit gegeben durch Veränderung der Welt von außen
- jetzt: zeitlicher Parameter explizit
- KB besteht aus Paar von Zeit t und Theorien ψ $\langle \Psi, t \rangle$ (temporale Spezifikation, hier nicht wichtig)
- keine Unterscheidung zwischen Revision und Update:
  - jetzt Tell(μ, t') mit $\langle \Psi, t \rangle$
  - sind die Zeitpunkte zu KB gleich t = t' dann gilt (ψ ○ μ, t)
  - wenn t' > t ist (ψ ◊ μ, t')
- Buch wird vom Roboter auf den Tisch gelegt, er meldet Aktion um 10.05am
  - Tell(b, 10.05am) --> es folgt Update
  - $\langle b, 10:05am \rangle$

Zeit, Revision und Update

- Luftaufnahme einer Situation um 10am, etwas liegt auf dem Tisch
- Spätere Analyse: es ist das Buch
- Tell(b, 10am) --> Revision (Zeit ist gleich)
- $\langle b \wedge \neg m, t \rangle$
- Jetzt ist es korrekt zu schließen das Magazin sei nicht auf dem Tisch, Revision ist intuitiv.
- Es ist ein Ziel viele Veränderungen in der Vergangenheit, Gegenwart und Zukunft auszuführen, daher reicht ein Paar von Theorie und Zeit nicht aus

Diskussion